THIS PAPER IS NOT TO BE REMOVED FROM THE EXAMINATION HALLS

UNIVERSITY OF LONDON **2910346**

BSc/Diploma Examination
for External Students

CREATIVE COMPUTING

Sound and Music: Sample Paper

Dateline:

Duration: 2 hours, 15 minutes

There are five questions in this paper. You should answer no more than **THREE** questions. Full marks will be awarded for complete answers to a total of **THREE** questions. Each question carries 25 marks. The marks for each part of a question are indicated at the end of the part in [.] brackets.

There are 75 marks available on this paper.

Electronic calculators must not be programmed prior to the examination. Calculators which display graphics, text or algebraic equations are not allowed.

Question 1 Computational Models of Music Cognition

(a) Within cognitive science, what is a computational theory, and what is its purpose? [6]

A computational theory is a formalisation of the computational requirements of a cognitive task, and is the highest level at which cognitive scientists typically study the mind. A cognitive theory is formulated by analysing the various outputs of a cognitive task resulting from different inputs. The specification of a computational theory guides experimental design, and the implementation of computer simulations, and provides a necessary explanatory framework in which to interpret empirical results.

(b) What do cognitive scientists mean by the *musical surface*? [4]

The musical surface is the level at which cognitive scientists study music. It is an abstraction of the physical musical signal into a sequence of discrete, musically salient events, such as notes and rests. It a methodological idealisation based on the psychological understanding of how music is perceived by the human mind, which allows scientists to study musical processes while not necessarily being concerned with the functionality of lower level perceptual mechanisms.

(c) The following table shows the symbolic representation of a melody in terms of basic attributes.

chromatic pitch	65	67	69	69	65	72	70	69	67	62	64	65
tonic reference	5	5	5	5	5	5	5	5	5	5	5	5
onset	0	4	8	10	12	20	22	24	30	32	35	36
duration	3	2	2	2	6	2	2	4	2	2	1	12

Extend this representation by writing down the values of the following two attributes:

i. chromatic interval [2]

ii. scale degree (note: tonic reference 5 is the key of F) [2]

chromatic interval	⊥	*2*	*2*	*0*	*-4*	*7*	*-2*	*-1*	*-2*	*-5*	*2*	*1*
scale degree	*0*	*2*	*4*	*4*	*0*	*7*	*5*	*4*	*2*	*9*	*11*	*0*

(d) In a symbolic music processing system, you should assume that the variables `chromaticInterval` and `onset` exist, each of which has been assigned an array containing the corresponding values in the table above and from your answer to question c.(i).

Write down a method for each of the following attributes, in Java code. The methods should accept a single array argument, and return a new array representing:

i. the contour of the melody; [4]

ii. the inter-onset-intervals of the melody. [4]

Example method for contour.

```
static int[] contour(int[] intervals)
{
  int[] contour = new int[intervals.length];
  for(int i=0; i<intervals.length; i++){
    int interval = intervals[i];
    if(interval == undefined){
      contour[i] = undefined;
    } else if (interval == 0){
      contour[i] = 0;
    } else if (interval > 0){
      contour[i] = 1;
    } else {
      contour[i] = -1;
    }
  }
  return(contour);
}
```

(e) Give an example of a learning model that has been published in the music cognition literature. Describe what and how the model is able to learn. [3]

The statistical model proposed withing the IDyOM project is a learning model developed by Marcus Pearce and other researchers based in the Intelligent Sound and Music Systems research group at Goldsmiths. The model learns the conditional probabilities of a range of musical surface melody attributes, using the Predication by Partial Match (PPM) algorithm. The model has a long-term model, which is trained on a corpus of music, representing a listener's musical knowledge, and a short-term model, which is trained on the melody being analysed. The model can be used to predict the likelihood of pitches given a melodic context, and by measuring uncertainty, the model can also be used to predict phrase boundaries.*

Question 2 Computer Music

(a) Name three early computer music pioneers, and for each, provide an example of why they are significant figures in the history of computer music. [6]

Example responses:

Pierre Schaeffer was a leader in the development of musique concrète, an approach to musical composition that began in the late 1940s, and utilised sound recordings as compositional material. He pioneered the use of early recording equipment for creative purposes, and also published theoretical writing on acousmatic music, which have had a considerable impact on later developments in electroacoustic and computer music aesthetics.

Iannis Xenakis was a composer and architect, who pioneered the use of mathematical models in compositional practice. He began composing seriously in the 1950, and published his influential book Formalized Music in 1971. Also in the 1970s, Xenakis developed the UPIC system, which allows the user to draw waveforms and amplitude envelopes as a means of generating new sounds and musical compositions.

(b) i. What is the distinction between ring modulation and classic amplitude modulation? [2]

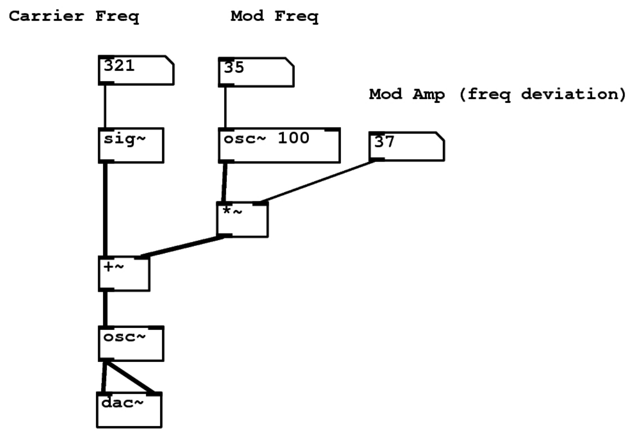
In amplitude modulation, the modulator signal is unipolar – i.e. it is a constant shifted signal typically in the range [0, 1]. This addition of DC means the frequency component of the carrier is present in the output (unlike ring modulation).

ii. Draw an accurately labelled diagram, in the frequency domain, of the signal resulting from the amplitude modulation of the following two signals: [6]

- carrier frequency $f_c = 1000$ Hz
- carrier amplitude $a_c = 1$
- modulator frequency $f_m = 360$ Hz
- modulator amplitude $a_m = 1$

The diagram should have the carrier frequency component at 1000 Hz, and contain two sidebands of 0.5 amplitude, at 640 Hz and 1360 Hz respectively. Additional mark for correct labelling.

(c) Draw a simple Pure Data patch producing the frequency modulation of two sinusoid signals, and sending the resulting signal to the computer's sound output. [5]



(d) Describe the following synthesis methods, and state the strengths and weaknesses of each technique from the point of view of a sound designer: [6]

i. subtractive synthesis;

ii. additive synthesis.

Subtractive synthesis is a method of generating new sounds by filtering harmonics from a source sound with a rich harmonic spectrum. Traditionally, oscillators such as pulse wave or sawtooth wave generators are used as sources, which create pitched sounds with strong overtones. Noise sources can also be used, which are particularly useful for synthesising percussive sounds. Such oscillators require only relatively simple analogue circuitry to build, or can be easily and efficiently implemented in software, so they offer a lot of flexibility to the sound designer at little cost. Also, the method is very intuitive, since the model of source + filter is the same as that of the human vocal tract. Although several oscillators can be combined, and low-frequency oscillators can be used to modify control signals in order to add variation, achieving very naturalistic or harmonically complex timbres can be difficult.

A similar description should also be provided for additive synthesis, which also makes similar observations about usability, practicality, and the sonic possibilities and limitations.

Question 3 Swarm music

(a) Write an essay discussing the use of principles of self-organisation in the generation of music. Consider both technical and aesthetic considerations, and also any potential limitations of the technique. [25]

The essay should contain a brief introduction, which summaries the argument and broader theme of the essay. There should also be a conclusion, again summarising key points, and offering some wider perspective. Intermediate paragraphs should each make a clearly defined statement, which together form a structured argument.

Some key points that might be addressed.

- *Self-organisation is a process often observed in the natural world.*
- *Self-organising systems are dynamic systems comprised of particles.*
- *Particles follow very simple reactive rules, there is no external or higher-level control.*
- *Complex coordinated behaviour emerges from a swarm of particles interacting within some environment.*
- *There are many approaches to generating music, some aim to explicitly model or learn aspects of musical style, others may be extremely abstract, or take inspiration from other domains (e.g. natural phenomena).*
- *Self-organisation may be an appropriate model for certain kinds of music - particularly improvised music.*
- *Swarm systems offer great interactive potential.*
- *There are problems in applying concepts from one domain (nature) to another (music).*
- *Practical decisions when using swarms to generate music. Using swarms to trigger notes? Or to generate separate pitch, rhythm, or dynamics data? Monophonic or polyphonic? Melodic or percussive?*
- *What kinds of tasks are swarms best suited for?*
- *Are swarms capable of generating sophisticated compositions? Or would they need combining with other generative techniques?*
- *Could swarms ever be genuinely creative?*

Question 4 Music information retrieval

(a) A digital music library typically contains a range of different kinds of digitised documents, ranging from audio recordings to scanned manuscripts.

List five different kinds of documents, each of which could be said to be an example of a particular piece of traditional folk music. [5]

Examples of digitised documents relating to a single piece of folk music might include:

- *field audio recordings of the work performed by traditional musicians;*
- *audio recordings of the work by other performers, for example, professional musicians in a recording studio, with possibly very different interpretations intended for specific audiences, for example, pop music releases;*
- *audio recordings of re-orchestrations, for example, of melodies used in classical or film music;*
- *audio recording which use samples of folk music, for example, remixes;*
- *scanned pages of traditional notations, for example canntaireachd;*
- *scanned pages of published anthologies of folk songs collected and notated by ethnomusicologists;*
- *scanned pages of re-orchestrations or piano reductions;*
- *computer files containing machine readable music notation;*
- *text files containing lyrics;*
- *MIDI file transcriptions.*

(b) Define the following evaluative criteria: [6]

i. precision;
ii. recall;
iii. F1.

i. Precision reflects the true positives as a proportion of the positive output of the model.

ii. Recall reflects the true positives as a proportion of the positives in the original data.

iii. F1 is the harmonic mean of precision and recall:

$$F1 = \frac{2 \cdot precision \cdot recall}{precision + recall}.$$

(c) In a music retrieval system containing audio recordings of pop songs and cover versions, describe two problems that will typically lead to poor recall performance. [6]

Two common problems would be global tempo variation (cover versions played at different speeds to the original), and also transposed variations (cover versions played in a different key). If both these issues were not addressed, recall performance would be low because the system would fail to recognise matches, resulting in a high number of false positives.

(d) Imagine that you have been hired by a record company to develop a system for identifying illegally used samples taken from recordings in their catalogue. Provide an overview design for such a system, and justify why the techniques you would use are suitable for this task. [8]

The answer to this question requires a sensible suggestion of infrastructure to enable the system to operate. For example, a database containing all of the record label's recordings must be accessible, and another database must be created containing all the extracted features from the record label's recordings. This second database should be automatically updated periodically as new material is added to the first database. If the system is to identify samples illegally used in recordings not owned by the company, a third database of features must be accessible or created, either in agreements with another record company, or from otherwise legally gathered audio.

Points to discuss.

- *An appropriate definition of 'similarity' needs to be given for this system.*
- *A range of appropriate features will be needed in order to capture such a potentially broad range matches (exact samples through to heavily transformed samples).*
- *A sequence of features might be more appropriate than a bag-of frames.*
- *Computational and efficiency issues - e.g. indexable features.*
- *The system is not operating under 'real time' constraints, so speed is not necessarily a major factor.*
- *How will the results be presented to the user in a useful way?*

Question 5 Music information retrieval

(a) Twelve-tone equal temperament is the tuning system typically used in Western classical and pop music.

Assuming the standard reference pitch A4 (MIDI note number 69) equals 440 Hz, calculate the frequency of the following pitches: [6]

i. A5

ii. D4

iii. A♯2

Requires calculating the MIDI note number for each pitch, and applying the following formula: $f(p) = 2^{\frac{p-69}{12}} \cdot 440$.

Pitch	*MIDI note number*	*Frequency (Hz)*
A5	*81*	*880.00*
D4	*62*	*293.66*
A♯2	*46*	*116.54*

(b) Explain the process of classical dynamic time warping. [6]

Complete answer should include the following points.

- *The objective of DTW is to find an optimal alignment between two time-dependent sequences (of length N and M).*
- *A local cost measure must be defined in order to compute the similarity between pairs of values.*
- *Compute a cost matrix (size* $N \times M$*) by applying the cost measure to each pair of elements of the sequences.*
- *Find an (N, M)- warping path with the minimal overall cost.*
- *A warping path must satisfy the boundary condition, monotonicity condition, and step size condition. These conditions should be briefly explained in words or mathematical notation.*
- *Describe the algorithm, based on dynamic programming, which recursively computes an accumulated cost matrix.*

(c) How might optical music recognition technology be combined with audio-based music information retrieval technology? Describe an application that could be built using a combination of both OMR and MIR technologies. [7]

OMR technology is able to take scanned images of musical scores, and generate symbolic representations of the musical notation. The quality of the generated data is variable, and errors can be frequent, particularly for low quality images or *non-standard scores. However, the availability of the symbolic information opens up possibilities for interesting applications.*

The answer to describe how the OMR generated symbolic data can be related to audio data (i.e. a music synchronisation task). The answer should also provide an interesting application, of genuine practical use, which combines scanned images of musical scores, and synchronised symbolic and audio music data.

(d) Describe the process of generating cepstrum features from an audio file. [6]

- *Apply a windowing function to convert the signal into frames.*
- *Calculate the discrete Fourier transform of each frame.*
- *Take the absolute values.*
- *Take the logarithm of the magnitudes.*
- *Apply inverse discrete Fourier transform.*

Additional mark for clarity and extra details.